# TOWARDS A NEURAL NETWORK MODEL OF THE ATTENTIONAL BLINK

H. BOWMAN AND B. P. WYBLE

*Computing Laboratory, University of Kent, Canterbury, Kent, CT2 7NF, UK*
*E-mail: {H.Bowman,B.Wyble}@kent.ac.uk*

P.J. BARNARD

*Cognition and Brains Sciences Unit, 15 Chaucer Rd, Cambridge, CB2 2EF, UK*
*E-mail: phil.barnard@mrc-cbu.cam.ac.uk*

One of the most prominent experimental paradigms for investigating the deployment of attention over time is the Attentional Blink (AB). Although there is now a great deal known about it, computational modeling of the AB remains only lightly explored. This paper responds to this limitation by proposing a prototype neural network model of the blink. A central aspect of which is a realization of the concept of consolidation into working memory, which is at the heart of the majority of current explanations of the blink.

## 1. Introduction

The majority of visual attention research has focused on the *spatial* dimension. Despite this traditional focus, there has been an increasing amount of research directed at the *temporal* profile of attention. This research has considered how long attention is "occupied" by performing a particular task. One prominent paradigm used to investigate this issue is the *Attentional Blink* (AB) [8]. There are now many variants of the AB paradigm, but one that can claim to be canonical, and is also the one we will focus on, locates two letter targets (which we denote T1 and T2) within an RSVP (Rapid Serial Visual Presentation) stream of digit distractors [2][a]. Items in the stream are presented at a rate of approximately 10 per second and the task is to identify the two targets in a report phase that follows the stream. The characteristic empirical finding is that report of the second target is poor if it appears within a certain time interval of the first, as typically demonstrated by a serial-position curve, such as that shown in figure 2a (basic blink condition).

---

[a] One reason for focusing on this experimental formulation is that it does not include a task switch, which has been argued to confound the blink paradigm [2].

Although there is now a great deal known about the AB, both in respect of empirical findings, e.g. [8,2] and proposed explanations, e.g. [10], computational modeling of the phenomenon remains only lightly explored. This paper responds to this limitation by proposing a prototype neural network model of the attentional blink. Our model is most naturally seen as a neural realization of Chun and Potter's two-stage explanation of the blink [2]. A central aspect of our model will be how it realizes the concept of consolidation into working memory, which is at the heart of Chun and Potter's and indeed the majority of current explanations of the blink. This model makes the further claim that the reason for closing the attentional gate is to allow accurate T1 binding, both in respect of binding together constituent features and binding to the correct temporal context.

The paper will begin by presenting background details on the blink in section 2. Then section 3 describes the model, section 4 presents the results of running the model and, finally, section 5 gives some concluding remarks.

## 2.  Background on the Blink and Theoretical Justification for Model

It would clearly be impossible to do justice to the spectrum of literature concerning the AB within the context of this paper. Thus, we will simply summarize the findings that most directly impinge upon our model. Firstly, the following are key characteristics of the AB serial position curve, see figure 2a (basic blink condition),

1.  the blink is a 100 to 500ms (approx) post T1 interval in which performance on T2 (conditional on correct T1 report) is significantly impaired;
2.  the blink generally has a sharper onset than offset;
3.  if T2 immediately follows T1 it is reported at baseline or near baseline levels (unless there is a substantial switch between T1 and T2 tasks); this is the *lag 1 sparing* phenomenon.

Modern explanations of the AB have been heavily influenced by studies that suggest that the blink has a late locus in the processing stream. Initial evidence for which came from priming studies, where it was found that with word based RSVP streams, missed T2 items primed a third target [11]. Furthermore, and perhaps even more compelling, evidence for the late locus hypothesis came from electrophysiological work, which showed that missed T2 items elicited electrical potentials associated with early perceptual activity (N1 and P2 waveforms) and with meaning (N400). However, working memory update waveforms (P3) were not present [13]. All of which suggests that the impairment to T2 processing occurs at the stage of consolidation into working

memory. That is, T2s that are not reported are, broadly speaking, processed as extensively at perceptual and semantic stages as T2s that are reported, however, consolidation of their perceptual and semantic traces into working memory is prevented.

These observations beg the question of what is meant by consolidation into working memory. The theoretical perspective that we will pursue is that a central element of consolidation is *binding*. That is, in order for an item to be successfully encoded into working memory its distributed neural representation needs to be bound into a coherent whole. In fact, there are two facets of binding that we will emphasize: (1) binding together the component features of an item and (2) binding items into the correct temporal context. Theories of binding have classically focused on the first of these. However, the second is also critical since in its absence it would be impossible to recall the order in which items were encoded into working memory, e.g. for subjects to know that T1 appeared before T2 in an AB stream. While correctly ordered recall is not a requisite for accuracy in most AB paradigms and experiments, Chun and Potter [2] recorded temporal order from subjects and found order to be conserved in the vast majority of reports at lags 2 and greater. Furthermore, we would argue that even in the absence of an explicit instruction to do so, there exists a default tendency to encode and recall the order of pairs of target items that are temporally dispersed.

This binding perspective offers one reason why the gate needs to be closed at the expense of T2 accuracy. Specifically, our working hypothesis is that T2 consolidation is suppressed in order to prevent interference with T1 binding. Thus, the blink is a mechanism to ensure coherent binding of T1s. According to this explanation, lag 1 sparing is a breakdown of the system; arising because the mechanism that ensures coherent T1 binding (by closing the gate on T2) is slow relative to the Stimulus Onset Asynchrony (SOA) used in the AB paradigm. A consequence of which is that binding errors should be observable at lag 1 because coherent binding of T1 is impaired by the T2, which enters into the binding process before the gate is closed. There is some support for this theory, since in many studies T1 performance is particularly poor at lag 1, as revealed, for example, by analysis of target report percentages in table 1 of [2]. However, it could be that this lag 1 binding breakdown shows up most significantly as T1 – T2 swaps, in which featural binding of both targets is (broadly) successful, but an erroneous temporal context binding has arisen, which makes it difficult for retrieval mechanisms to correctly identify the temporal order of T1 and T2 at lag-1. Chun and Potter [2] (see figure 8 on page 119) demonstrate exactly this sort of error as do preliminary experiments conducted within in our lab. One

reason for constructing the model that we present here is to investigate whether this coherent binding interpretation is consistent with the available AB data.

A final empirical finding that will greatly influence us is the observation that placing a blank after either of the targets (i.e. at the T1+1 or T2+1 positions) greatly attenuates and even in some cases eradicates the blink [8,2]. This suggests that targets are backward masked by the immediately following items [9]. That is, since they appear in the same spatial location, the iconic traces of stream items compete at a preattentive stage, with the trace elicited by a target being curtailed by the arrival of the immediately following item.

Our neural network model has also been strongly influenced by the two-stage explanation of the blink proposed by Chun and Potter [2]. In their first stage all stimuli are processed to a preliminary level at which features and perhaps even meaning are extracted. However, this level is subject to rapid forgetting and is not sufficient for report. It is only through stage 2 that stimuli are consolidated to a level required for a response. In contrast to stage 1, this second stage is capacity limited and thus, creates a bottleneck at which T2's decay while T1 is being processed

## 3. The Model

Our neural network model is depicted schematically in figure 1. The main layers are an *input layer* at which items are presented; a *masking layer* at which preattentive visual traces of items compete; a *category layer* at which task relevant items are foregrounded; and finally, a *working memory* mechanism, through which items are encoded and retrieved. We discuss each of these in turn. Except for the two Working Memory layers, all layers of the model use representations that are localist in nature. Thus, these layers contain one neuron for each type of item that can appear in the RSVP stream. Future extensions of the model will provide for distributed representations, but for our purposes here, simple localist representations are sufficient for testing our hypotheses.

**Masking Layer.** RSVP items are presented in sequence at the input layer, which feeds activation forward to the masking layer, where feedback inhibition forces these activity traces to compete. It is through this mechanism that backward (and in fact, forward) masking effects are realized. For example, in the absence of any further input, a strongly active neuron in the layer will slowly decay back to zero. However, the trace of an active neuron would be rapidly curtailed if a second item arrived at the layer during this decay period.

**Category Layer.** The masking layer feeds activation forward to the *category* layer. A task demand unit selectively foregrounds neurons that code

target items (e.g. letters) and suppresses those that code background items (e.g. digits). Items in this layer do not compete, but their initial activation reflects the activity of stimulus traces in the masking layer. In particular, items that are masked will yield *weak traces* at the category layer and those that are unmasked will yield *strong traces*.
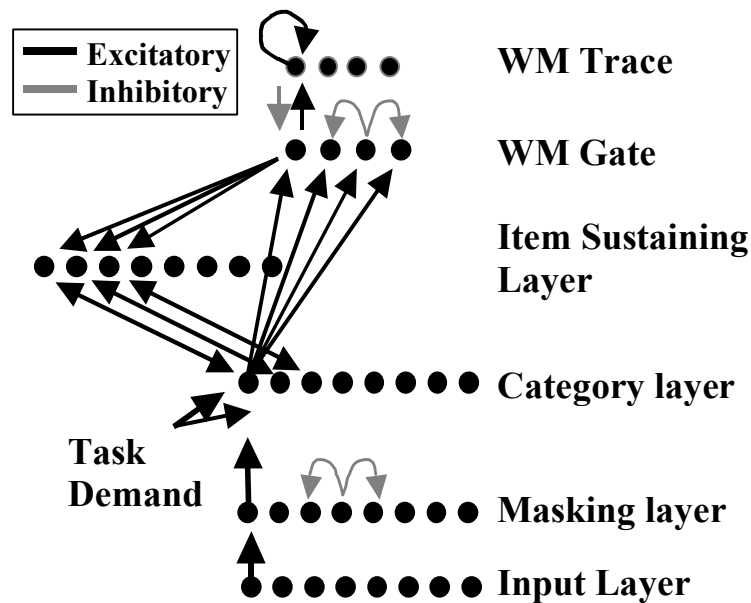


Figure 1. The Full Model. Note that the Hebbian binding links from the WM Gates to the Category layer are not included, as they play no role in the functional dynamics, serving only to indicate when successful binding has occurred.

**Working Memory.** As previously stated the theory that we are exploring explains the blink in terms of binding targets into Working Memory (WM). These ideas are implemented through interaction between the WM and the localist representations in the category layer. The category layer and working memory layers interact through three mechanisms. First, each category neuron directly excites all of the WM gates. Second, direct, hebbian binding links are established between specific WM gates and category layer neurons. Finally, at least one WM gate neuron has to be available to activate the item sustaining layer, which allows the system to encode, and later recall, weak (masked) items.

A pair of neurons, consisting of a WM gate and a WM trace together serve as a token [7]. Tokens are used to indicate that a target was identified, what that item was, and in what order it was perceived relative to other targets in the

stream. Thus, a WM token is an encoding that combines what and when, effectively creating an "instance" of a "type". The WM gate neurons control access to their respective WM trace neurons, allowing or preventing category layer items from activating that trace. Strong winner-take-all dynamics within the gate neuron layer ensure that only one trace neuron is accessible at a time. Each trace neuron is individually self-excitatory and can self-sustain for the duration of a trial once activated. Working memory consolidation is implemented in terms of building *binding links* between WM gates and category layer neurons. Thus, an item is viewed as having been consolidated into working memory if a link is successfully built from a WM gate to the neuron coding that item at the category layer. Conceptually, these links should be viewed as pointers from a given WM *token* to the featural / categorical neural circuits that code the *type* of the item being consolidated, which enable later top-down retrieval of that item. Binding links are built via a rapid hebbian process between WM gates and category layer neurons[b]. These links are unidirectional and play no part in the functional dynamics of the model during the presentation of the input.

In addition to WM neurons serving as pointers in order to enable later retrieval, they also code temporal order. This is obtained by, firstly, using a winner-take-all mechanism amongst WM gates to ensure that only one gate neuron is active at any instant; this neuron denotes the current temporal context. Furthermore, each WM trace neuron sends an inhibitory projection to its corresponding gate neuron. This ensures that when a WM trace neuron has been activated, it closes its own WM gate for the remainder of the trial. The closing of one such gate initiates winner-take-all competition amongst the remaining WM gate neurons until a new gate is made available for future binding.

Thus, gates are made available in sequence in an order determined by the bias inputs applied to them. In "normal" processing, i.e. when the time gap between pairs of (to be consolidated) targets is long, each WM gate obtains a link to a single category layer item, denoting that a different temporal context has been allocated to each recognized target. However, when target items arrive at the category layer in close temporal proximity (as arises at short lags in AB streams), the handover between WM gates can be too slow to keep up. Consequently, binding errors can occur, which arise in the model when a single WM gate obtains links to multiple category layer items. In addition, T2s can be missed altogether because their category layer activation falls in the window

---

[b] This use of Hebbian learning might be viewed as controversial, however, it is not essential to our model and could be replaced by an activation-based gating mechanism.

between one WM gate (which has encoded T1) being suppressed and the next one becoming available.

**Item Sustaining Layer.** The final mechanism that we need to explain is the *item sustaining layer*. As previously discussed, masking plays an important role in obtaining the blink. This is reflected in our model since binding links can only be constructed between strongly active WM gates and strongly active category neurons (this is built into our Hebbian learning rule). Thus, in the absence of further stimulation, binding fails for category layer traces of masked targets.

Conceptually, we believe that the brain provides a mechanism to "grab hold" of such weak, but *task relevant*, stimuli and sustain them for a sufficiently long period that they can be bound into working memory. The item sustaining layer implements such a mechanism by providing a recurrent excitatory circuit to prolong the duration of traces in the category layer neuron in order that they can be encoded (similar techniques can be found elsewhere, e.g. [4]). For this excitatory circuit between a given category neuron and its dedicated sustaining neuron to be active, concurrent input from both the category layer and a WM gate must be present. Since each WM gate connects to all of the sustaining neurons, any one active gate will enable all of the sustaining neurons. As the network completes the process of encoding the T1, the system undergoes a switch from one active WM token to the next. During this handoff process, all of the WM gates are inactive and consequently all of the sustaining neurons follow suit. It is during this handoff from one token to the next that the system is no longer able to encode masked items because of the inactivity of the item sustaining layer. It is imperative that the item sustaining layer is temporarily shut down in this way or the first item of any sequence would be encoded to all of the available WM tokens. The subject would recall multiple instances of an item that was presented only once. This scenario is strongly contra indicated by data on the repetition blindness effect [7].

The length of time required by this system to encode an item is determined by the strength of that item. Strong (unmasked) items are rapidly encoded by strongly activating WM gates, which causes binding links to be built quickly. Conversely, weak items bind more slowly and require prolonged assistance from the item sustaining neurons.

This is the heart of how the model blinks, yet exhibits lag 1 sparing. A T2 presented immediately after a T1 has a chance of building a binding link with the same WM gate neuron. While this is technically an error, we propose that the system is able to disambiguate this double-encoding during retrieval in a process not explicitly modeled here. We further propose that it is because two

items are bound to the same token that it is difficult to correctly recall the order of these items. This has been shown by experimental work of Chun and Potter [2] who demonstrate a selective impairment in the recall of temporal order at lag 1.

If the SOA between a masked (and therefore weak) T1 and T2 is between 200 and 400 ms, the WM door will close prior to creation of the binding link to the T2 and the category trace will fade and be lost during the switch from one WM gate to the next. In this way the door is closed on T2 while T1 is being bound into WM and the length of time it takes T1 to be bound regulates the length of the blink. Unmasked (and therefore strong) T1's are able to establish binding links with the appropriate WM neuron more rapidly. Thus, the blink is shorter and shallower when T1s are unmasked rather than masked.

A MATLAB implementation of this model is available at www.cs.kent.ac.uk/people/staff/bw5/ncpwblinkmodel/.

## 4.  Results

The crucial performance measure of the Attentional Blink paradigm is the successful encoding of T2 for trials in which T1 was encoded (T2|T1). In this model, encoding of a target was scored as successful if the binding links from any WM gate to that target were above a designated threshold. Thus, we have not at this stage considered how the number of T1 – T2 swaps varies with lag, although this information could easily be extracted from our model. The elements of the conventional AB paradigm that will be modeled explicitly in this paper include the deficit at lags 2-5, the relative sparing of performance at lag 1, and the attenuation of the blink curve by blank(s) following T1 and T2.
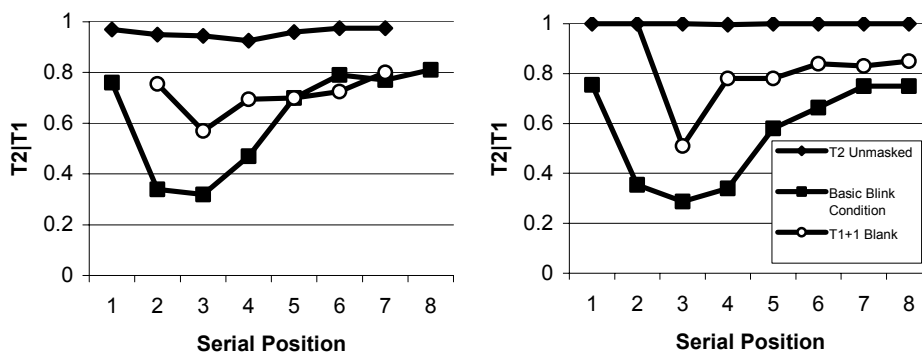


Figure 2. (a) Human data (on left) and (b) model data (on right).

The data being modeled will be extracted from [2] and [6]. Both of these papers used paradigms that are generally equivalent to the one used here. SOA was 100ms and targets were digits in a letter stream. T1 and T2 tasks were identical. This data is available from the aforementioned papers for the three conditions studied: basic blink condition [2], T1+1 blank [2] and T2 as the final element of the stream [6]. Figure 2a presents these three conditions for the experimental data while Figure 2b presents data for the same three conditions in the model. It should be clear from these results that, at least in qualitative terms, our model successfully reproduces these three experimental conditions.

The elevated performance for lag 2 in the T1 +1 Blank condition is due to a lack of forward masking for the T2 item, which increases its strength, compounding the attenuated blink caused by the stronger T1 trace. The experimental data from Chun and Potter [2] indicates the same sort of effect at lag 2, although to a lesser degree. In our model, any manipulations that increase the strength of T1 and T2 items attenuates the blink. Therefore this forward masking effect, inherent in the design of the masking layer, allows our model to make the prediction that the attentional blink will be attenuated by blanks at positions T1-1 and T2-1.

## 5.  Conclusions

We have presented a prototype neural network model of the AB, which has enabled us to explore how key AB data can be reproduced by a WM consolidation model. According to this theory, in order to protect T1 from binding errors, a door is closed on the consolidation of T2 targets if their category layer trace falls within the window of T1 binding. This is the blink window. However, the door is not closed instantaneously. A result of which is that lag 1 T2s can be consolidated. However, this process can erroneously bind T1 and T2 to the same WM token. This error allows T1 and T2 to be retrieved at lag 1, but without their correct temporal order. In addition, the model reproduces AB masking effects. Specifically, the blink is attenuated if either the T1+1 or the T2+1 items are left blank. In the former case this arises because strong T1s are rapidly consolidated into working memory and thus, a fresh WM token is released before T2 has decayed, while in the latter case the blink is attenuated because stronger T2s can out-live the blink window.

It is beyond the scope of this paper to give a detailed comparison of our model to existing theories of working memory and prefrontal function. However, it is safe to say that our model has similarities to a number of such

theories, e.g. in respect of foregrounding task relevant items, c.f. [3,1]; allocation of general purpose WM resources (our WM gate and trace neurons), c.f. Duncan's adaptive coding theory [5]; sustaining activation by setting up reverberating circuits, c.f. Dehaene et al's global workspace resource [4]; and active maintenance which occurs in our WM trace neurons, c.f. [12].

## Acknowledgements

## References

1. Braver, T.S. and J.D. Cohen (2000) "On the Control of Control: The Role of Dopamine in Regulating Prefrontal Function and Working Memory" Control of Cognitive Processes, Attention and Performance XVIII, S. Monsell and J. Driver (eds), MIT Press, 713-737.
2. Chun, M. M. and M. C. Potter (1995). "A Two-Stage Model for Multiple Target Detection in Rapid Serial Visual Presentation." J Exp Psych: HPP **21**(1): 109-127.
3. Cohen, J.D., K. Dunbar, et al. (1990) "On the Control of Automatic Processes: A Parallel Distributed Processing Account of the Stroop Effect" Psychological Review **97**(3): 332-361.
4. Dehaene, S., M. Kerszberg et al (1998) "A Neuronal Model of a Global Workspace in Effortful Cognitive Tasks" Proc Natl Acad Sci USA, **95**: 14529-34.
5. Duncan, J. (2001) "An Adaptive Coding Model of Neural Function in Prefrontal Cortex" Nature Reviews, Neuroscience, **2**, Nov., 820-829.
6. Giesbrecht, B. and V. Di Lollo (1998) "Beyond the Attentional Blink: Visual Masking by Object Substitution" J Exp Psych: HPP **24**(5): 1454-66.
7. Kanwisher N. (1991) "Repetition blindness and illusory conjunctions: errors in binding visual types with visual tokens." J Exp Psych: HPP **17**(2):404-21.
8. Raymond, J.E., K.L. Shapiro et al (1992) "Temporary Suppression of Visual Processing in an RSVP Task: An Attentional Blink" J Exp Psych: HPP 18(3):849-860.
9. Seiffert, A.E. and V. Di Lollo (1997) "Low-Level Masking in the Attentional Blink" J Exp Psych: HPP **23**(4): 1061-1073.
10. Shapiro, K. L., K. M. Arnell, et al. (1997). "The Attentional Blink." Trends in Cognitive Sciences **1**(8): 291-296.
11. Shapiro, K. L., J. Driver, et al. (1997). "Priming from the attentional blink: A failure to extract visual tokens but not visual types." Psych Sci **8**: 95-100.

12. Usher, M. and Davelaar, E. (2003) "Short/long term memory in terms of activation versus weight based processes" this volume.

13. Vogel, E.K., S.J. Luck et al (1998). "Electrophysiological Evidence for a Postperceptual Locus of Suppression During the Attentional Blink" <u>J Exp Psych: HPP</u> **24**: 1656-1674.